



The  
Patent  
Office

PC- 4837/01158

09/171671

PROPERTY DOCUMENT

The Patent Office  
Concept House  
Cardiff Road  
Newport  
South Wales  
NP9 1RH

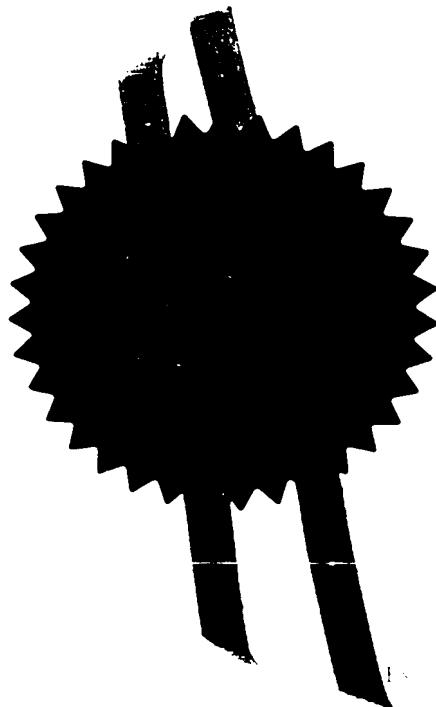
REC'D 26 MAY 1997  
WPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed

Dated 14th May 1997



The  
Patent  
Office

24 APR 1996

1649824 E199117-2 (0157) 2001 0700 15.00

The Patent Office

Cardiff Road  
Newport  
Gwent  
NP9 1RH

## Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form.)

Fee: £25

1. Your reference

37520/JMD

2. Patent application number

(The Patent Office will fill in this part)

9608457.9

3. Full name, address and postcode of the or of each applicant (*underline all surnames*)PEPTIDE THERAPEUTICS LIMITED,  
321, Cambridge Science Park,  
Milton Road,  
Cambridge CB4 4WG.Patents ADP number (*if you know it*)

If the applicant is a corporate body, give the country/state of its incorporation

4. Title of the invention

INHIBITORS OF PROTEOLYTIC ENZYMES

5. Full name, address and postcode in the United Kingdom to which all correspondence relating to this form and translation should be sent

Reddie & Grose  
16 Theobalds Road  
LONDON  
WC1X 8PLPatents ADP number (*if you know it*)

91001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (*if you know it*) the or each application number

Country

Priority application number  
(*if you know it*)Date of filing  
(day month year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing  
(day month year)

*if any applicant name in part 5 is not an inventor  
or  
if there is an inventor who is not named as an  
applicant, or  
if the application is not filed by a person  
described in part 5*

## Patents Form 1/77

Enter the number of sheets for any of the following items you are filing with this form.  
Do not count copies of the same document

Continuation sheets of this form	0
Description	14
Claim(s)	0
Abstract	0
Drawing(s)	8



10 If you are also filing any of the following, state how many against each item.

Priority documents	0
Translations of priority documents	0
Statement of inventorship and right to grant of a patent ( <i>Patents Form 7/77</i> )	0
Request for preliminary examination and search ( <i>Patents Form 9/77</i> )	0
Request for substantive examination ( <i>Patents Form 10/77</i> )	0
Any other documents (please specify)	0

11

We request the grant of a patent on the basis of this application.

Signature

Date



24 April 1996

12 Name and daytime telephone number of person to contact in the United Kingdom

DR. J.M. DAVIES  
0171 - 242 0901

### Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or such direction has been revoked.

### Notes

- a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.
- b) Write your answers in capital letters using black ink or you may type them.
- c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- e) Once you have filled in the form you must remember to sign and date it.
- f) For details of the fee and ways to pay please contact the Patent Office.

### Inhibitors of Proteolytic Enzymes

The present invention relates to the field of inhibitors of proteolytic enzymes and to methods for identifying active substrates for proteolytic enzymes.

Many therapeutically useful drugs act as enzyme inhibitors. In particular, proteolytic enzyme inhibitors have been the focus of much attention in the pharmaceutical industry; because they play a variety of roles in a multitude of biological systems. Their proteolytic activities are related to processes ranging from cell invasion associated with metastatic cancer to evasion of an immune response, as seen in certain parasitic organisms; from nutrition to intracellular signalling to the site-specific proteolysis of viral proteases and eukaryotic hormone-processing enzymes. However, the traditional random screening methods for the identification of lead molecules as inhibitors of proteolytic enzymes are often laborious and time-consuming. Therefore new and efficient methods which can accelerate the drug discovery process are greatly in demand.

We consider that proteases contain an active catalytic site which tends to become increasingly activated as the recognition pockets ( $S_1$  and  $S_2$  etc) and ( $S_{1'}$  and  $S_{2'}$  etc) become better occupied. Therefore, it is important that those parts ( $P_1$  and  $P_2$  etc) ( $P_{1'}$  and  $P_{2'}$  etc) of the inhibitor that best fit into these pockets are identified as quickly as possible in order to design novel protease inhibitors. Therefore, we have devised a combinatorial method for the rapid identification of these binding motifs which will greatly expedite the synthesis of inhibitors of a variety of proteolytic enzymes such as aspartyl proteases, serine proteases, metallo proteases and cysteiny1 proteases.

The use of a fluorescence resonance energy substrate (FRET) for the analysis of proteolytic enzyme specificity was first published by Carmel.<sup>1</sup> Since then many different quenched fluorogenic substrates for measuring enzyme inhibition have been described in the literature.<sup>2-11</sup> These substrates contain a fluorophore, F, in a P position (*vide supra*), which is quenched by another group, Q, present in a  $P'$  position (*vide supra*) and separated from F by the scissile bond. The advantage of the positioning of these substrates, F and Q, is that cleavage of a peptide bond occurs between the two substrates and, therefore, represents a more natural hydrolytic event rather than the cleavage and release of a C-terminal chromophore.

<sup>1</sup> Carmel, M. *J. Am. Chem. Soc.* 1962, 84, 10200. ABz = N-(4-nitrophenyl)-4-nitroanilide moiety. Upon hydrolysis of the amideacyl-4-nitroanilide bond, the highly fluorescent N-ABz group is released attached either to an amino acid or peptide.

Immobilised libraries, where substrates are attached to a polymer or biopolymer support, have also been used for mapping protease binding sites.<sup>13</sup> Singh et al reported recently that enzymatic substrate activity of 38 selected octapeptides attached via a linker to controlled pore glass is predictive of the same activity of similar peptides in solution. However, these results are preliminary and only for a specific example. Therefore, it is not clear whether immobilised substrates attached to polymers can reliably replace soluble substrates in mapping the hindered protease binding sites especially since the hydrophilic or lipophilic nature of the polymer and the size of the interstices within the polymer are bound to influence the reaction between the enzyme and its substrates.

Mixtures of internally quenched, fluorogenic substrates have also recently been described in which the quencher group, Q, is 2,4-dinitrophenyl (Dnp) and is attached to the P side of the scissile bond, while the fluorogenic group, is N-methyl anthranilic acid (Nma) and is attached to the P' side.<sup>14</sup>

The specificity of soluble peptide libraries have been determined.<sup>15,16</sup> Berman et al. described<sup>16</sup> an HPLC mass spectrometry technique in which 6 mixtures of 128 peptides were synthesised which were N-terminally labelled with the Dnp group in order to allow UV monitoring on the HPLC. The disadvantage of this approach is that each assay mixture has to be individually analysed, because no fluorogenic substrate is revealed, and that the effective concentration of each separate component is limited by the size of the mixture because of overall solubility factors.

Drevin et al.<sup>17</sup> have suggested the use of individually synthesised fluorogenic substrates for the determination of enzyme activity using a chromophore which chelates lanthanide ions. Garman and Phillips have suggested the use of FRET substrates in which the fluorogenic and quencher moieties are attached via thiol or amino functional groups after the peptide has been synthesised, but this has the disadvantage that they are not in library form and that these functional amino and thiol groups need to be selectively revealed after the peptide has been synthesised. Wang et al. have suggested the use of the EDANS and DABCYL fluorescor and quencher pairing for the individual synthesis of substrates for proteolytic enzymes.

The above methods which have used FRET techniques for the mapping of the active site around a specific protease suffer from one or more of the following disadvantages:

- i. because of general aqueous insolubility they do not produce mixtures of compounds in a form suitable for high throughput screening in aqueous solution
- ii. the derivatised compounds cannot be prepared in combinatorial library form using solid phase techniques.
- iii. the mixtures which have been used<sup>8,9</sup> were not self-decoding, and needed time-consuming deconvolutive resynthesis for identification of the active molecules

## Brief Description of the Invention

The present invention relates to the field of:

- i. the inhibition of proteolytic enzymes.
- ii. the rapid generation of structure-activity relationships using auto-deconvoluting combinatorial libraries, which facilitate the invention of novel inhibitors of proteolytic enzymes.
- iii. the detection and measurement of proteolytic enzyme activity using combinatorial FRET (fluorescence energy transfer) libraries of molecules.
- iv. the establishment of biological assays for proteolytic enzymes through the rapid discovery of highly active substrates for proteolytic enzymes

We describe herein a process which can be used for the rapid generation of structure-activity relationship (SAR) data and, therefore, the characterisation of the binding motif of any protease, e.g. an endoprotease, and which will, therefore, facilitate:

- i. the development of a sensitive enzyme inhibition assay by using the best compound in the library as the fluorogenic substrate for the proteolytic enzyme under scrutiny,
- ii. the invention of novel proteolytic enzyme inhibitors by rapid characterisation of the best binding motif,
- iii. computer aided drug design to design potent inhibitors using known methodology, and also in prioritising which pre-synthesised compounds in the in-house and commercially available databases to assay.

Two complementary compound libraries, L1 and L2, each containing  $n \times 1600$  compounds, of the type A-B<sub>1-10</sub>-C<sub>1-10</sub>-D<sub>1-8</sub>-n(E<sub>1-2</sub>)-F-G, are prepared in which

A = a fluorescor internally quenched by F, preferably an unsubstituted anthranilic acid derivative, connected by an amide bond to B.

B, C, D, E, are natural or unnatural amino acid residues connected together by amide bonds, although B, C, D and E can be any set of groups, provided that the scissile bond between D-E is an unsubstituted amide bond.

F = a quencher capable of internally quenching the fluorescor A, preferably an unsubstituted or substituted 3-nitrotyrosine derivative.

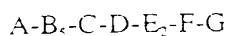
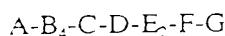
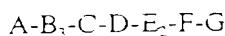
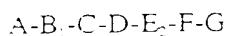
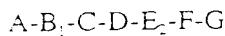
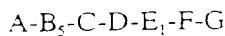
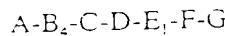
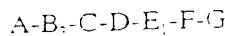
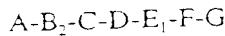
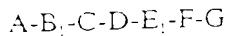
G = optionally present and is a hydrophilic moiety, preferably as aspartyl amide moiety.

If present, G advantageously ensures that all compounds in the library are imparted with

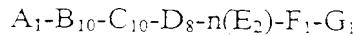
solubility and/or water solubility for any type of assay.

Note that A and F herein correspond generally and respectively to moieties B and Q of the prior art referred to above).

The numbers represented in subscript following residues B, C, D and E refer to the number of possibilities from which those residues are selected. Thus, by way of illustrative example, A-B<sub>1,5</sub>-C-D-E<sub>1,2</sub>-F-G represents a mixture of the following ten compounds:



The general combinatorial formula for the library can be expressed as:



providing  $1 \times 10 \times 10 \times 8 \times n \times 2 \times 1 \times 1 = 1600n$  compounds.

Both compound libraries, L1 and L2, of the above type are synthesized using solid phase techniques using the Multipin approach<sup>24</sup> such that each library contains 1600n compounds as 80n mixtures of 20 distinct, identifiable compounds. These 20 component mixtures are then placed separately into each of 80 wells of a 96 well plate (the other two lanes are used for control experiments) and then screened against a known quantity of the protease.

Thus it is an important part of the invention that regardless of the number of compounds contained in the two libraries L1 and L2 (1600n, where n = any integer between 1 and 4) the libraries themselves are complementary and amenable to deconvolution without recourse to resynthesis. It is also an important part of the invention that the library matrix has been especially formatted so that the most important site pairings P<sub>2</sub> and P<sub>1</sub> for proteolytic enzymes can be identified immediately without recourse to resynthesis.

Those compounds of the type A-B-C-D-E-F-G, that are the better substrates for the protease will be cleaved, and can be readily identified because the fluorescor, A, will be cleaved from its nearby quencher F, in a time dependent manner which can be easily quantified. The fluorescent quenching by F of A only occurs when the two are in nearby proximity, normally within 20 angstrom units. Hence cleavage of the bond D-E allows F to move further away from A and thus allow A to fluoresce when excited by light of the correct wavelength.

In this manner the most active compound can be rapidly identified without the need for further resynthesis and deconvolution. Moreover, the wells that show the most rapid development of fluorescence can also be analysed by mass spectrometry, since by comparison with the original mixture, the identity of the most efficient substrate can be found by its disappearance into its two component parts, e.g A-B-C-D and E-F-G. Hence the problem of library deconvolution can be overcome and the most active substrate for the enzyme can be rapidly identified.

In addition, after the initial treatment of the proteolytic enzyme with the library mixtures, L1 and L2, the residual enzymatic activity in each well can be quantified by the addition of a the most potent fluorogenic substrate for the enzyme, which is found in the 6400 compound library. Because of the nature of the library design this can be quickly prepared and purified. If there is no appearance of increased fluorescence with the known substrate, S1, then the presence of an enzyme inhibitor can be inferred, which again can be quickly identified without the need for resynthesis.

The general description of the library layout will now be described with reference to figures 1 to 14.

For example, when  $n = 1$  and the library contains 1600 compounds, in the first column of the first row (A1) (Fig. 1) in the first plate (P1) of the library L1, (hereinafter designated as location A1,P1,L1) there will be one C component, C<sub>1</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>1</sub> and E<sub>2</sub>) (Fig. 2). In the tenth column of the first row (A10) in the first plate (P1) of the library L1, (hereinafter designated as location A10,P1,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>1</sub> and E<sub>2</sub>). In the tenth column of the eighth row (H8) in the first plate (P1) of the library L1, (hereinafter designated as location H8,P1,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>8</sub>, the ten B components and the two E components (E<sub>1</sub> and E<sub>2</sub>). Hence all 1600 components are present in the one plate, because the 80 wells each contain 20 components.

A second complementary library is synthesised as follows (Fig. 3). In the first column of the first row (A1) of the first plate (P1) of the library, L2, (hereinafter designated as location A1,P1,L2), there will be ten C components, two D components (D<sub>1</sub> and D<sub>2</sub>), one B component, B<sub>1</sub>, and one E component, E<sub>1</sub>. In the tenth column of the first row (A10) of the first plate (P1) of the library, L2, (hereinafter designated as location A10,P1,L2), there will be ten C components, two D components (D<sub>1</sub> and D<sub>2</sub>), one B component, B<sub>1</sub>, and one E component, E<sub>1</sub>. In the first column of the second row (B1) of the first plate (P1) of the library, L2, (hereinafter designated as location A10,P1,L2) there will be ten C components, two D components (D<sub>1</sub> and D<sub>2</sub>), one B component, B<sub>1</sub>, and one E component, E<sub>2</sub>. In the tenth column of the second row (B10) of the first plate (P1) of the library, L2, (hereinafter designated as location A10,P1,L2) there will be ten C components, two D components (D<sub>1</sub> and D<sub>2</sub>), one B component, B<sub>10</sub>, and one E component, E<sub>2</sub>. Hence only the first two rows are used to accommodate 400 compounds in total.

In the first column of the first row (A1) of the second plate (P2) of the library, L2 (hereinafter designated as location A1,P2,L2), there will be ten C components, two D components (C<sub>3</sub> and C<sub>4</sub>), one B component, B<sub>1</sub>, and one E component, E<sub>1</sub> (Fig. 4). In the tenth column of the first row (A10) of the first plate (P2) of the library, L2 (hereinafter designated as location A10,P2,L2), there will be ten C components, two D components (D<sub>3</sub> and D<sub>4</sub>), one B component, B<sub>10</sub>, and one E component, E<sub>1</sub>. In the first column of the second row (B1) of the second plate (P2) of the library, L2, (hereinafter designated as location B1,P2,L2), there will be ten C components, two D components (D<sub>3</sub> and D<sub>4</sub>), one B component, B<sub>1</sub>, and one E component, E<sub>2</sub>. In the tenth column of the second row (B10) of the second plate (P2) of the library, L2, (B10,P2,L2), there will be ten C components, two D components (D<sub>3</sub> and D<sub>4</sub>), one B component, B<sub>10</sub>, and one E component, E<sub>2</sub>. Hence only the first two rows are used to accommodate 400 compounds in total.

In the first column of the first row (A1) of the third plate (P3) of the library, L2, (hereinafter designated as location A1,P3,L2), there will be ten C components, two D components (D<sub>5</sub> and D<sub>6</sub>), one B component, B<sub>1</sub>, and one E component, E<sub>1</sub> (Fig. 5). In the tenth column of the first row (A10) of the third plate (P3) of the library, L2, (hereinafter designated as location A10,P3,L2), there will be ten C components, two D components (D<sub>5</sub> and D<sub>6</sub>), one B component, B<sub>10</sub>, and one E component, E<sub>1</sub>. In the first column of the second row (B1) of the third plate (P3) of the library, L2, (hereinafter designated as location B1,P3,L2), there will be ten C components, two D components (D<sub>5</sub> and D<sub>6</sub>), one B component, B<sub>1</sub>, and one E component, E<sub>2</sub>. In the tenth column of

the second row (B10) of the third plate (P3) of the library, L2, (B10,P3,L2), there will be ten C components, two D components (D<sub>5</sub> and D<sub>6</sub>), one B component, D<sub>10</sub>, and one E component, E<sub>2</sub>. Hence only the first two rows are used to accomodate 400 compounds in total.

In the first column of the first row (A1) of the fourth plate (P4) of the library, L2, (hereinafter designated as location A1,P4,L2), there will be ten C components, two D components (D<sub>7</sub> and D<sub>8</sub>), one B component, B<sub>1</sub>, and one E component, E<sub>1</sub> (Fig. 0). In the tenth column of the first row (A10) of the fourth plate (P4) of the library, L2, (hereinafter designated as location A10,P4,L2), there will be ten C components, two D components (D<sub>7</sub> and D<sub>8</sub>), one B component, B<sub>10</sub>, and one E component, E<sub>1</sub>. In the first column of the second row (B1) of the fourth plate (P4) of the library, L2, (hereinafter designated as location B1,P4,L2), there will be ten C components, two D components (D<sub>7</sub> and D<sub>8</sub>), one B component, B<sub>1</sub>, and one E component, E<sub>2</sub>. In the tenth column of the second row (B10) of the fourth plate (P4) of the library, L2, (B10,P4,L2), there will be ten C components, two D components (D<sub>7</sub> and D<sub>8</sub>), one B component, D<sub>10</sub>, and one E component, E<sub>2</sub>. Hence only the first two rows are used to accomodate 400 compounds in total.

In this fashion two complementary libraries, L1 and L2 are prepared. In library, L1, each of the 80 of wells contains a mixture of 20 components providing 1600 compounds for screening. In library, L2, four plates are used in which only the first two columns and rows are employed, providing 20 wells of 20 components per well per plate, and furnishing the same 1600 compounds in library, L1, but in a format in which no two compounds found together in library, L1, will be found together in library, L2.

Thus it is an important part of the invention that the compounds contained in the two libraries L1 and L2 are novel and that the libraries themselves are complementary, in that any two compounds which are found together in a 20 component mixture in the same location (e.g. A1P1L1) in library, L1, are not found together in any of the 20 component mixtures in any location of the library, L2.

In analogous examples, where separately n=2, 3 or 4, extra plates are constructed in library, L1, format to accomodate the component pairs E<sub>1</sub> and E<sub>2</sub> (n=2), E<sub>1</sub> and E<sub>3</sub> (n=3), and E<sub>1</sub> and E<sub>4</sub> (n=4), respectively. For the respective deconvolution libraries of the type, L2, the respective rows in the plates P1, P2, P3, and P4, are increasingly filled with the paired components D<sub>1</sub> and D<sub>2</sub>, D<sub>1</sub> and D<sub>3</sub>, and D<sub>1</sub> and D<sub>4</sub>, respectively.

For example, when  $n = 3$ , and the library contains 4800 compounds, in the first column of the first row (A1) in the first plate (P1) of the library L1, (hereinafter designated as location A1,P1,L1) there will be one C component, C<sub>1</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>1</sub> and E<sub>2</sub>). In the tenth column of the first row (A10) in the first plate (P1) of the library L1, (hereinafter designated as location A10,P1,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>1</sub> and E<sub>2</sub>). In the tenth column of the eighth row (H8) in the first plate (P1) of the library L1, (hereinafter designated as location H8,P1,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>8</sub>, the ten B components and the two E components (E<sub>1</sub> and E<sub>2</sub>). Hence 1600 components are present in the one plate, because the 80 wells each contain 20 components.

In the first column of the first row (A1) in the second plate (P2) of the library L1, (hereinafter designated as location A1,P2,L1) there will be one C component, C<sub>1</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>3</sub> and E<sub>4</sub>). In the tenth column of the first row (A10) in the second plate (P2) of the library L1, (hereinafter designated as location A10,P2,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>3</sub> and E<sub>4</sub>). In the tenth column of the eighth row (H8) in the second plate (P2) of the library L1, (hereinafter designated as location H8,P2,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>8</sub>, the ten B components and the two E components (E<sub>3</sub> and E<sub>4</sub>). Hence 1600 components are present in the one plate, because the 80 wells each contain 20 components.

In the first column of the first row (A1) in the third plate (P3) of the library L1, (hereinafter designated as location A1,P3,L1) there will be one C component, C<sub>1</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>5</sub> and E<sub>6</sub>). In the tenth column of the first row (A10) in the third plate (P3) of the library L1, (hereinafter designated as location A10,P3,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>5</sub> and E<sub>6</sub>). In the tenth column of the eighth row (H8) in the third plate (P3) of the library L1, (hereinafter designated as location H8,P3,L1) there will be one C component, C<sub>10</sub>, one C component, C<sub>8</sub>, the ten B components and the two E components (E<sub>5</sub> and E<sub>6</sub>). Hence 1600 components are present in the one plate, because the 80 wells each contain 20 components. In total the three plate, P1, P2 and P3, contain 1600 compounds.

For example, when  $n = 4$ , and the library contains 6400 compounds, in the first column of the first row (A1) in the first plate (P1) of the library L1, (hereinafter designated as location A1,P1,L1) there will be one C component, C<sub>1</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>1</sub> and E<sub>2</sub>) (Fig. 7). In the tenth column of the first row (A10) in the first plate (P1) of the library L1, (hereinafter designated as location A10,P1,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>1</sub> and E<sub>2</sub>). In the tenth column of the eighth row (H8) in the first plate (P1) of the library L1, (hereinafter designated as location H8,P1,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>8</sub>, the ten B components and the two E components (E<sub>1</sub> and E<sub>2</sub>). Hence all 1600 components are present in the one plate, because the 80 wells each contain 20 components.

In the first column of the first row (A1) in the second plate (P2) of the library L1, (hereinafter designated as location A1,P2,L1) there will be one C component, C<sub>1</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>3</sub> and E<sub>4</sub>) (Fig. 8). In the tenth column of the first row (A10) in the second plate (P2) of the library L1, (hereinafter designated as location A10,P2,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>3</sub> and E<sub>4</sub>). In the tenth column of the eighth row (H8) in the second plate (P2) of the library L1, (hereinafter designated as location H8,P2,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>8</sub>, the ten B components and the two E components (E<sub>3</sub> and E<sub>4</sub>).

In the first column of the first row (A1) in the third plate (P3) of the library L1, (hereinafter designated as location A1,P3,L1) there will be one C component, C<sub>1</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>5</sub> and E<sub>6</sub>) (Fig. 9). In the tenth column of the first row (A10) in the third plate (P3) of the library L1, (hereinafter designated as location A10,P3,L1) there will be one C component, C<sub>10</sub>, one D component, C<sub>1</sub>, the ten B components and the two E components (E<sub>5</sub> and E<sub>6</sub>). In the tenth column of the eighth row (H8) in the third plate (P3) of the library L1, (hereinafter designated as location H8,P3,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>8</sub>, the ten B components and the two E components (E<sub>5</sub> and E<sub>6</sub>).

In the first column of the first row (A1) in the fourth plate (P4) of the library L1, (hereinafter designated as location A1,P4,L1) there will be one C component, C<sub>1</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>7</sub> and E<sub>8</sub>) (Fig. 10). Likewise, in the tenth column of the first row (A10) in the fourth plate (P4) of the library L1, (hereinafter designated as location A10,P4,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>1</sub>, the ten B components and the two E components (E<sub>7</sub> and E<sub>8</sub>). In the tenth column of the eighth row (H8) in the fourth plate (P4) of the library L1, (hereinafter designated as location H8,P4,L1) there will be one C component, C<sub>10</sub>, one D component, D<sub>8</sub>, the ten B components and the two E components (E<sub>7</sub> and E<sub>8</sub>).

one B component, B<sub>1</sub>, and one n component, E<sub>1</sub> (Fig. 11). In the tenth column of the first row (A10) of the first plate (P1) of the library, L2, (hereinafter designated as location A10,P1,L2), there will be the ten C components, two D components (D<sub>1</sub> and

$D_3$ ), one B component,  $B_{10}$ , and one E component,  $E_1$ . In the first column of the eighth row (H1) of the first plate (P1) of the library, L2, (hereinafter designated as location H1,P1,L2), there will be the ten C components, two D components ( $D_1$  and  $D_2$ ), one B component,  $B_1$ , and one E component,  $E_8$ . In the tenth column of the eighth row (H10) of the first plate (P1) of the library, L2, (H10,P1,L2) there will be the ten C components, two D components ( $D_1$  and  $D_2$ ), one B component,  $B_{10}$ , and one E component,  $E_8$ . Hence the matrix containing all ten columns and all eight rows are used to accomodate 1600 compounds in total.

In the first column of the first row (A1) of the second plate (P2) of the library, L2, (hereinafter designated as location A1,P2,L2), there will be ten C components, two D components ( $D_3$  and  $D_4$ ), one B component,  $B_1$ , and one E component,  $E_1$  (Fig. 12). In the tenth column of the first row (A10) of the first plate (P2) of the library, L2, (hereinafter designated as location A10,P2,L2), there will be ten C components, two D components ( $D_3$  and  $D_4$ ), one B component,  $B_{10}$ , and one E component,  $E_1$ . In the first column of the second row (B1) of the second plate (P2) of the library, L2, (hereinafter designated as location B1,P2,L2), there will be ten C components, two D components ( $D_3$  and  $D_4$ ), one B component,  $B_1$ , and one E component,  $E_2$ . In the tenth column of the eighth row (H10) of the second plate (P2) of the library, L2, (H10,P2,L2), there will be ten C components, two D components ( $D_3$  and  $D_4$ ), one B component,  $B_{10}$ , and one E component,  $E_8$ .

In the first column of the first row (A1) of the third plate (P3) of the library, L2, (hereinafter designated as location A1,P3,L2), there will be ten C components, two D components ( $D_5$  and  $D_6$ ), one B component,  $B_1$ , and one E component,  $E_1$  (Fig. 13). In the tenth column of the first row (A10) of the third plate (P3) of the library, L2, (hereinafter designated as location A10,P3,L2), there will be ten C components, two D components ( $D_5$  and  $D_6$ ), one B component,  $B_{10}$ , and one E component,  $E_1$ . In the first column of the second row (B1) of the third plate (P3) of the library, L2, (hereinafter designated as location B1,P3,L2), there will be ten C components, two D components ( $D_5$  and  $D_6$ ), one B component,  $B_1$ , and one E component,  $E_2$ . In the tenth column of the eighth row (H10) of the third plate (P3) of the library, L2, (H10,P3,L2), there will be ten C components, two D components ( $D_5$  and  $D_6$ ), one B component,  $B_{10}$ , and one E component,  $E_8$ .

In the first column of the first row (A1) of the fourth plate (P4) of the library, L2, (hereinafter designated as location A1,P4,L2), there will be ten C components, two D components ( $D_7$  and  $D_8$ ), one B component,  $B_1$ , and one E component,  $E_1$  (Fig. 14). In the tenth column of the first row (A10) of the fourth plate (P4) of the library, L2, (hereinafter designated as location A10,P4,L2), there will be ten C components, two D components ( $D_7$  and  $D_8$ ), one B component,  $B_{10}$ , and one E component,  $E_1$ . In the first column of the second row (B1) of the fourth plate (P4) of the library, L2, (hereinafter designated as location B1,P4,L2), there will be ten C components, two D components ( $D_7$  and  $D_8$ ), one B component,  $B_1$ , and one E component,  $E_2$ . In the tenth column of the eighth row (H10) of the fourth plate (P4) of the library, L2, (H10,P4,L2), there will be ten C components, two D components ( $D_7$  and  $D_8$ ), one B component,  $B_{10}$ , and one E component,  $E_8$ .

Abbreviations used herein

are as follows: Abz, aminobenzoyl; DMF, dimethylformamide; Fmoc, fluorenylmethoxycarbonyl; HPLC, high performance liquid chromatography; Pfp, pentafluorophenyl, tBoc, tert-butoxycarbonyl; tBu, tert-butyl; TFA, trifluoroacetic acid; Pmc, pentamethyl chroman, Pbf, pentamethylbenzofuran, TBTU, 2-(1H-Benzotriazole-1-yl)-1,1,1,3,3-tetramethyluronium tetrafluoroborate; HOBr, N-hydroxybenzotriazole; Trt, Trityl.

The compounds were synthesised using the Multipin approach<sup>22,23</sup> using Fmoc-Rink amide Macro crowns (Chiron Mimotypes Pty., Ltd., 11 Duerdin Street, Clayton Victoria 3168, Australia) with a loading of 7  $\mu$ Moles. The coupling chemistry employed is similar to that reported in the literature<sup>27</sup> for fluorenylmethoxycarbonyl protected amino acids and activated pentafluorophenyl esters, in which the side-chains are protected using acid labile protecting groups known to those skilled in the art, such as Boc- (for the -NH<sub>2</sub> of Lysine, and -NH<sub>2</sub> of anthranilic acid), tBu- (for the -OH groups of serine, threonine and tyrosine), t-But for the -COOH group of Aspartic acid and Glutamic acid, Trityl- (for the Amide of Asparagine and Glutamine, and the amine functionality of the Histidine ring; Pmc or Pbf (for the guanidino function of arginine). The N- $\alpha$ -fluorenylmethoxycarbonyl protecting group of the coupled residues were cleaved using 20% piperidine in dimethylformamide (DMF) for 30 minutes at 20° C. The coupling reactions for the free acids such as Boc-ABz-OH (Boc-2-aminobenzoic acid), and Fmoc-(3-nitro)tyrosine-OH were accomplished using 10 equivalents of a mixture of the free acid (1 eq.) :TBTU (0.98 eq.); HOBr (0.98 eq.) : N-methylmorpholine (1.96 eq.) in dimethylformamide (500  $\mu$ L) as solvent for 2 hours at 20° C. The other amino acids were coupled as their pentafluorophenyl esters<sup>25</sup> for 2 hours.

Hence, in order to couple approximately equal ratios of each component in the mixture of the derivatised amino acids as their pentafluorophenyl esters, a solution of a total of 0.95 equivalents (relative to the amino group loading on the resin) of the mixture of amino acid pentafluorophenyl esters HOBT (1 eq.) in DMF (500  $\mu$ L) were coupled for 16 hours at 20° C. The pins were then washed well with DMF and then recoupled using the same mixture under the same conditions. Using a 16 hour double coupling protocol with equimolar mixtures of the derivatised pentafluorophenyl esters of the amino acids in slightly less than 1 equivalent, it is possible to obtain approximately equal amounts of the coupled products to the pin. In this fashion the libraries are constructed with 20<sup>3</sup> compounds present on each crown of resin. The compounds were cleaved from the crowns directly into the 80 designated wells of the desired 96 well plate (Fig. 1). In the cleavage protocol each crown was treated with a mixture (500  $\mu$ L) containing trifluoroacetic acid (95%), triethylsilane (5%) for 2 hours at 20° C. The crowns were then washed with trifluoroacetic acid (500  $\mu$ L) and this was then combined with the cleavage solution.

The Fmoc-Rink amide Macro crowns (Chiron Mimotypes Pty. Ltd., 11 Duerdin Street, Clayton Victoria 3168, Australia) at 7  $\mu$ Mol loading per crown, were coupled

After 16 hours in DMF for 16 minutes, and subsequent washing with DMF and then methanol, coupling of the Fmoc-(3-nitro)tyrosine-OH was accomplished using 10

equivalents of a mixture of the Fmoc-(3-nitro)tyrosine-OH (1 eq.) : TBTU (0.98 eq.) : HOBt (0.98 eq.) : N-methylmorpholine (1.96 eq.) in dimethylformamide as solvent at 0.14 M concentration for 2 hours at 20° C. Removal of the Fmoc group (*vide infra*) was followed by coupling of the mixtures of amino acids in the ratios outlined and under the conditions described (*vide infra*).

In a particular example the amino acids comprising group B include Ala, Val, Leu, Ser, Asn, Gln, Glu, Lys, Phe, Pro. The amino acids comprising group C include Ala, Val, Leu, Ser, Asn, Gln, Glu, Lys, Phe, Pro. The amino acids comprising group D include Ala, Val, Ile, Leu, Nle, Ser, Glu, Phe. For n=4, the amino acids comprising group E include Ala, Val, Ile, Leu, Nle, Ser, Glu, Phe. Otherwise any selection from the amino acids can be made for n=1, 2, or 3.

The plates containing the combined cleavage solutions were then evaporated to dryness to yield the component mixtures using a rotary centrifuge ("SPEEDVAC", Savant Instruments Inc., Farmingdale, NY) at 800 rpm for 1 hour at 20° C under a reduced pressure of 10<sup>-2</sup> mmHg. Each component was then transferred to the final mother plate (Fig. 1) using a (50%: 45%: 5%) mixture of acetonitrile: water: acetic acid. The plates were then lyophilised to dryness using at 20° C under a reduced pressure of 10<sup>-2</sup> mmHg, and then stored under nitrogen at -20° C. In this fashion libraries of the type shown in Figures 2-14 were prepared.

The novel methodology described herein greatly facilitates the invention of therapeutically useful proteolytic enzyme inhibitors and is commercially exploitable. This is because the best substrate motif for the proteolytic enzyme can be rapidly identified, and, since there exists in the literature a variety of ways for attaching motifs which react with the active site of a proteolytic enzyme, especially for aspartyl, metallo, serine and cysteinyl proteases, an enzyme inhibitor can be readily synthesised. Moreover, amide bond replacements or transition state mimetics which can be incorporated into the molecule, and would be especially useful for the inhibition of aspartyl or metallo proteases.

The method described also facilitates the rapid development of a screening assay for novel protease inhibitors. The most potent fluorogenic substrate discovered by library screening can subsequently be used for the detection of inhibitors of the particular proteolytic enzyme under scrutiny.

The presence of an inhibitor within the compound libraries described is readily detected by retreatment of the assay mixture with the most active fluorogenic substrate, which will allow the immediate measurement of the remaining proteolytic enzyme activity.

The invention provides self-decoding, combinatorial fluorogenic libraries, and it will greatly facilitate the design and invention of novel protease inhibitors because:

- i. The peptides of the library may have increased aqueous solubility in comparison to peptides containing similar and other fluorogenic and quencher groups.
- ii. The peptides are stable to contaminating exopeptidases.
- iii. The self deconvolution method described, coupled with the continuous analysis of the rate of substrate cleavage data, allows the immediate identification of the most active binding motif contained within the substrate library.
- iv. The method allows for the rapid assessment of the enzyme assay mixture for any compounds in the library that are acting as enzyme inhibitors.

## References.

1. I. Schlechter and A. Berger, Biochem. Biophys. Res. Commun., 1967, 27, 157-162)
2. A. Carmel et al., FEBS Lett., 1973, 30, 11.
3. M. M. Meldal and I. Svendsen, J. Chem. Soc. Perkin Trans.1, 1995, 1591-1596,
4. M. Meldal and K. Breddam, Anal. Biochem., 1991, 195, 141-147,
5. T. Forster, Ann. Phys., 1948, 6, 55.
6. A. Yaron, A. Carmel, and E. Katchalski-Katzir, Anal. Biochem. 1979, 95, 228 and references therein.
7. S.A. Latt et al., Anal. Biochem., 1972, 50, 56.
8. A. Persson et al., Anal. Biochem., 1977, 83, 2,96.
9. I. Yu Filippova et al., Bioorg. Khim., 1986, 12, 1172.
10. J. Pohl et al., Anal. Biochem., 1987, 165, 96.
11. S.J. Pollack et al., J. Am. Chem. Soc., 1989, 111, 5961.
12. E. K. Bratovanova and D.D. Petkov, Analytical Biochem., 1987, 162, 213.
13. J. Singh et al., J. Med. Chem., 1995, 38, 217-219 and references therein.
14. M. Green et al. in 'Innovation and Perspectives in Solid Phase Synthesis' (R Epton Ed.) Mayflower Worldwide Ltd., Birmingham, UK. 1994, 239-244.
15. J. R. Petithory et al., Proc. Natl. Acad. Sci. USA, 1991, 88, 11510-11514
16. J. Berman et al., J. Biol. Chem., 1992, 267, 1434-1437.
17. H. Drevin, A. -T. Martin, J. Carlsson, S. Oscarsson, T. Lovgren, I. Hemmila and M. Kwiatkowski, WO 89/10975, (5-5-1988)
18. A. J. Garman and N. G. Phillips, WO 94/28166 (27-5-1993)
19. G.T. Wang and E.D. Matayoshi, E.P. 428000 (3-11-1989)
20. G. A. Krafft, G. T. Wang and E.D. Matayositi, EP 428000, (3-11-1989).
21. G. R. Marshall and M. V. Toth, USP 5,164,300, (11-12-1990).
22. G. R. Marshall and M. V. Toth, USP 5,011,910, (28-12-1989).
23. K. T. Chapman, N. A. Thornberry, M. MacCoss, J. R. Weidner, R. A. Mumford, W. K. Hagmann, EP 528487A (16-8-1991)
24. R. P. Haugland, WO 93/04077, (23-8-1991)
25. R. M. Valerio, A. M. Bray N. J. Maeji, Int. J. Peptide Protein Res., 1994, 44, 158-165
26. M. Bastos, N. J. Maeji and R. H. Abeles, Proc. natl. Acad. Sci., 1995, 92, 6738-6742.
27. 'Solid Phase Peptide Synthesis', E. Atherton and R. C. Sheppard, IRL Press 1989.

	1	2	3	4	5	6	7	8	9	10	11	12
A	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10		
B	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10		
C	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10		
D	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10		
E	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10		
F	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10		
G	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10		
H	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10		

Figure 1

2/8

First Example of Library Matrix where n=1

Component Distribution in Plate 1, Library 1 (n=1)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
D1	B <sub>1-10</sub> E <sub>1-2</sub>									
D2	B <sub>1-10</sub> E <sub>1-2</sub>									
D3	B <sub>1-10</sub> E <sub>1-2</sub>									
D4	B <sub>1-10</sub> E <sub>1-2</sub>									
D5	B <sub>1-10</sub> E <sub>1-2</sub>									
D6	B <sub>1-10</sub> E <sub>1-2</sub>									
D7	B <sub>1-10</sub> E <sub>1-2</sub>									
D8	B <sub>1-10</sub> E <sub>1-2</sub>									

Figure 2

Component Location in Plate 1, Library 2 (n=1).

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
E1	C <sub>1-10</sub> D <sub>1-2</sub>									
E2	C <sub>1-10</sub> D <sub>1-2</sub>									

Figure 3

## Component Location in Plate 2, Library 2 (n=1).

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
E1	C <sub>1-10</sub> D <sub>3-4</sub>									
E2	C <sub>1-10</sub> D <sub>3-4</sub>									

Figure 4

## Component Location in Plate 3, Library 2 (n=1).

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
E1	C <sub>1-10</sub> D <sub>5-6</sub>									
E2	C <sub>1-10</sub> D <sub>5-6</sub>									

Figure 5

4/8

### Component Location in Plate 4, Library 2 (n=1).

**Figure 6**

#### Example Library where n=4

### Component Distribution in Plate 1, Library 1 (n=4)

**Figure 7**

## Component Location in Plate 2, Library 1 (n=4)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
D1	B <sub>1-10</sub> E <sub>3-4</sub>									
D2	B <sub>1-10</sub> E <sub>3-4</sub>									
D3	B <sub>1-10</sub> E <sub>3-4</sub>									
D4	B <sub>1-10</sub> E <sub>3-4</sub>									
D5	B <sub>1-10</sub> E <sub>3-4</sub>									
D6	B <sub>1-10</sub> E <sub>3-4</sub>									
D7	B <sub>1-10</sub> E <sub>3-4</sub>									
D8	B <sub>1-10</sub> E <sub>3-4</sub>									

Figure 8

## Component Location in Plate 3, Library 1 (n=4)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
D1	B <sub>1-10</sub> E <sub>5-6</sub>									
D2	B <sub>1-10</sub> E <sub>5-6</sub>									
D3	B <sub>1-10</sub> E <sub>5-6</sub>									
D4	B <sub>1-10</sub> E <sub>5-6</sub>									
D5	B <sub>1-10</sub> E <sub>5-6</sub>									
D6	B <sub>1-10</sub> E <sub>5-6</sub>									
D7	B <sub>1-10</sub> E <sub>5-6</sub>									
D8	B <sub>1-10</sub> E <sub>5-6</sub>									

6/8

Component Location in Plate 4, Library 1 (n=4).

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
D1	B <sub>1-10</sub> E <sub>7-8</sub>									
D2	B <sub>1-10</sub> E <sub>7-8</sub>									
D3	B <sub>1-10</sub> E <sub>7-8</sub>									
D4	B <sub>1-10</sub> E <sub>7-8</sub>									
D5	B <sub>1-10</sub> E <sub>7-8</sub>									
D6	B <sub>1-10</sub> E <sub>7-8</sub>									
D7	B <sub>1-10</sub> E <sub>7-8</sub>									
D8	B <sub>1-10</sub> E <sub>7-8</sub>									

Figure 10

Component Location in Plate 1, Library 2 (n=4).

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
E1	C <sub>1-10</sub> D <sub>1-2</sub>									
E2	C <sub>1-10</sub> D <sub>1-2</sub>									
E3	C <sub>1-10</sub> D <sub>1-2</sub>									
E4	C <sub>1-10</sub> D <sub>1-2</sub>									
E5	C <sub>1-10</sub> D <sub>1-2</sub>									
E6	C <sub>1-10</sub> D <sub>1-2</sub>									
E7	C <sub>1-10</sub> D <sub>1-2</sub>									
E8	C <sub>1-10</sub> D <sub>1-2</sub>									

Figure 11.

718

#### Component Location in Plate 2, Library 2 (n=4).

Figure 12

### Component Location in Plate 3, Library 2 (n=4).

Figure 13

8/8

## Component Location in Plate 4, Library 2 (n=4).

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
E1	C <sub>1-10</sub> D <sub>7-8</sub>									
E2	C <sub>1-10</sub> D <sub>7-8</sub>									
E3	C <sub>1-10</sub> D <sub>7-8</sub>									
E4	C <sub>1-10</sub> D <sub>7-8</sub>									
E5	C <sub>1-10</sub> D <sub>7-8</sub>									
E6	C <sub>1-10</sub> D <sub>7-8</sub>									
E7	C <sub>1-10</sub> D <sub>7-8</sub>									
E8	C <sub>1-10</sub> D <sub>7-8</sub>									

Figure 14.

